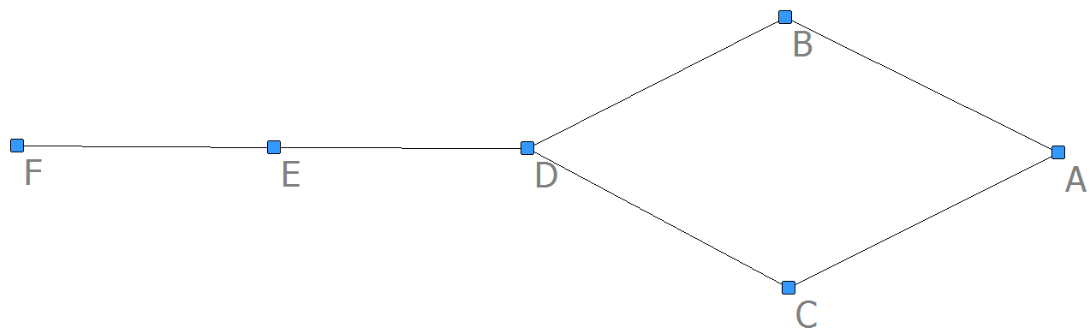
Sociology 2275, Social Network Analysis, Fall 2022

Assignment 4: Centrality Measures

This assignment provides some experience with basic centrality concepts and common centrality measures available in the xUCINET software.

A. First, consider the following small undirected network.



Do the following by hand:

1. Calculate the degree centrality of all vertices, both unnormalized and normalized.
2. Calculate the betweenness centrality for vertices B and E, both unnormalized and normalized.
3. Calculate the unnormalized eccentricity for vertices D and A.
4. Calculate the unnormalized closeness centrality for vertex C.

B. Do this part of the assignment using the data project on the study about the karate club that underwent a schism, `Zachary_KarateClub`, that is stored as part of the xUCINET package. Study the binary version of the network, `Zachary_KarateClub$Connection`, not the valued one `Zachary_KarateClub$Strength`.

1. Obtain degree centrality scores for the club members. What four are most central? Which two are least central?

2. Visualize the club member network, letting node sizes correspond to degree centrality and node colors correspond to post-schism club membership (see attributes stored as part of data project). Turn in a copy of the plot you construct.

[NB: gplot does not automatically rescale vertex size; when keying size to some property (like an attribute value or centrality score), the node size will depend on the scale of the attribute or score. The default vertex size in gplot is 1, so it may help to rescale your degree values via multiplication or addition, such that they average about 1 when you set the `vertex.cex` argument that governs vertex size in the gplot command. You might accomplish this by dividing the degrees by their mean.

If your vector of degrees is stored in a vector v_degree, first calculate its mean as

```
>mean_deg<-mean(v_degree[,])
```

Then use this when you designate vertex size within the gplot function

```
vertex.cex=v_degree/mean_deg
```

You may need to experiment some in order to obtain a display with visually appealing vertex sizes. There is also a useful user-defined “rescale” function on pp. 65-66 of Douglas Luke’s book, *A User’s Guide to Network Analysis in R*.]

3. Obtain betweenness centrality scores for the club members. Do betweenness and degree order them in markedly different ways?

4. Obtain eigenvector (reflection) centralities for the network, and use them in a graph to highlight the most important club members (see the above on scaling vertex sizes). Compare this graph to the one from part 2. Can you identify vertices that appear more important when their proximity to other prominent vertices is taken into account (do this by comparing degree and eigenvector centralities)?

5. Obtain Freeman closeness centrality scores for the club members. Then assess the extent to which degree, closeness, betweenness, and eigenvector/reflection scores converge for this network by correlating them.

[First, construct an R “data frame” that includes the different scores, as follows:

```
>cent_scores<-data.frame(<degree scores>[,1],
     <closeness scores[,1],
     <betweenness scores>[,1],
     <eigenvector scores>[,1]
     )
> colnames(cent_scores)<-c("Degree",
     "Closeness","Betweenness","Eigenvector")
```

Then use the correlation function `cor()` in base R to obtain the Pearson correlation coefficients among all pairs of scores.

```
>cor(cent_scores)
```
]

Which pair of scores is most closely associated? Which pair of them is least closely linked?

6. Calculate and compare normalized centralization values for this network, for the degree and betweenness measures (see software notes about centrality for how to accomplish this). For which type of centrality is the inequality among club members evidently greatest?

Due: Thursday, October 6.

Note: These assignments are not graded, but we do take note of whether or not you do them.
Please submit to either derick_baum@g.harvard.edu or pvm@wjh.harvard.edu .